DOCUMENT RESUME

ED 091 422 TM 003 636

AUTHOR Warries, Egbert
TITLE Standard Mastery Curves and Skew Curves.
PUB DATE [Apr 74]
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (59th, Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Educational Research; Evaluation Needs; *Graphs; *Measurement Techniques; *Scores; Statistical Analysis; *Summative Evaluation; Testing; Test Interpretation
IDENTIFIERS Mastery Learning; Skew Curves; *Standard Mastery Curves

ABSTRACT

The objective of the study is to convince educational researchers of the necessity for "standard mastery curves" for the graphical representation of scores on summative tests for a group of students. Attention is drawn to the study of theoretical and empirical skew curves in education and biology. Use of standard mastery curves and study of skew curves in statistics and biology is in consequence with the basic idea of mastery learning. It seems to fill the gap between felt pressure to reach mastery and the relative measurement devices available. (Author)

26.02

WMW.90/74/148
11th April 1974.

SUMMARY

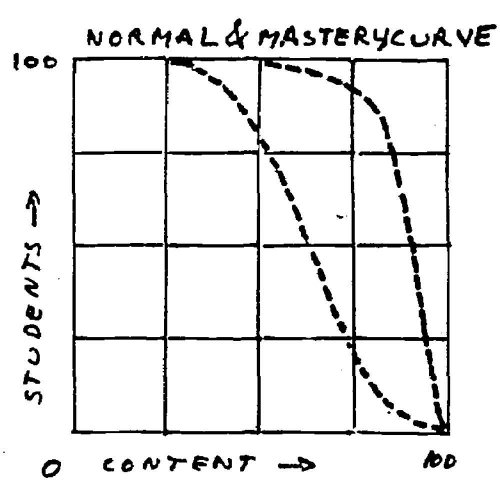
STANDARD MASTERY CURVES AND SKEW CURVES

AERA APRIL 19TH, 1974.

Egbert Warries
Twente University of Technology
P.O. Box 217, Enschede, Holland

1. The main objective of the paper is to indicate the necessity of the use of other graphical presentations of summative student scores when testing for mastery is concerned. Furthermore descriptions for "non-mastery strategies" are sought to explain the scape of curves of poor mastery. In a final note the attention is asked for the possibility of theorizing about skew curves and the necessity in education to give more attention to other curves than the normal one.

2. All kind of solutions have been offered to get rid of the drawbacks of relative measurement and the psychological mechamisms active when a person observes a normal curve. Nothing does really help as long as we record the outcomes of teaching in such a way that the differences in student performance hit you in the eye. Does specifying objectives of instruction in behavioral terms help? Does developing teaching and learning strategies in which more students than before reach a criterion help? Relative measurement is a psychological necessity.



3. The frequency table of raw scores and the frequency curve are no longer usable , when our main interest is in recording the fact that student x reached a point on the attainment scale. The registration that I recommend, is based on the wish to provide an answer to the questions: how many students reached the score x? and which score did y percent of my students reach? The matching graphical presentation is a curve in a square of 100 x 100 millimeters, in which the horizontal line represents the content of the course or the summative test, (1% = 1 mm), reaching a certain point on the horizontal axis. The recommended method is to draw the line which connects all the possible dots in answer to the question: how many reached this score?

4. The method recommended for theorizing about standard mastery curves in education is based on assumptions about the causes of drop out of students during the course. Dividing the course in a number of tasks, we can assume many different values for the mastery/nonmastery ratio after each task, the probability to master a task after missing one or more preceeding ones and the relation of task-mastery to summative test.

5. The data to be used in the paper to illustrate this mode of representing scores on summative tests in a mastery setting, are both empirical and theoretical. In the empirical curves use is made of testresults at eleven-plus examination in Amsterdam and course-examination in Twente University of Technology. In the theoretical curves we have assumed that the summative test is a representation of each of a series of learning tasks of which the course consists.

6. Concluding remarks: The use of a standard mastery curve seems to be in consequence with the basic ideas of the mastery learning strategy. The further direct advantages seem to be: the attention is drawn to classmastery not to individual differences; obligation to discuss criterion levels and objectives of the course; clear illustration of surface in the square to be covered yet.
   Also in consequence with mastery learning seems to be a further study of the causes that make skew curves; both theory and description of reality in education and in natural science could be of use.

26.02

Planning Papers III

WMW.90/74/147
11th April 1974

Standard Mastery Curves and Skew Curves*

Egbert Warries,
Twente University of Technology,
The Netherlands.

1 In the last few years the work with which educational researchers occupy themselves seems, more than in earlier years, to indicate that schools as a system become more retentive and less apt to continuing selection. Both instruction and achievement testing follow this tendency, the former by developing instructional methods which garantee succes for more students than formerly and the latter by developing testing methods which try to measure the performance of the students in a more absolute manner than our norm-referenced methods do.

In this paper, I occupy myself with the measurement of learning outcomes in school after a considerable period of instruction, perhaps 100 hours. When schools really no longer are in need of a scoring system in which the student scores almost by law are normally distributed, then it is my thesis that this actually means that we are no longer interested in differentiation of student scores. If this is true, and I hold it to be so, we will have to use other registrations of learning outcomes of individuals and groups. Also we have to use other graphical representations of group results. In this paper I propose such other graphical presentations, which I did call standard mastery curves. It is my objective, to convince the reader of the necessity of other performance-registrations and graphical representations and to indicate the need for study of the statistical characteristics of mastery curves like the characteristics of the normal curve.

*Prepared for Annual Meeting, American Educational Research Association, Chicago, april 1974.

2 The need for scoring methods more directly related to the content of what is learned, led Ebel (1962) to his notion of content scores, in contrast to normative standard scores which are based on the relative achievement performances of those who have taken the test. The unfitness of relative measurement in the evaluation of instructional methods was stressed by Glaser (1963) who made the distinction between criterion referenced and norm referenced measurements. As indicated in both articles, when achievement testing is used to provide information about individual behavior, the primary emphasis always is to discriminate among individuals and seldom or not to give information about the position of the individual on a proficiency scale.

3 Of course all kinds of solutions have been offered to get rid of the drawbacks of relative measurement and the mechanisms that start working when we observe a normal curve. Well known are the efforts to equate scores of different years and different examinations. There are also efforts to measure more consistently by some kind of grading on the curve; in our country we tried this through a marking system based on the standard error of measurement (Wijnen 1971) and by a system in which the group achievement on a number of "kernel"-items in the test, sets the criterion for the whole test (De Groot 1964). Nedelsky (1954) offered the system of the hypothetical borderline-person and Ebel (1962) offered the content standard test scoring system. For my examinations I sometimes make use of naive subjects to set the zero-line, which is also done by others.
However, all this does not help as long as we record the outcomes of our teachings in such a manner that the differences between the student performances hit you in the eye directly, (Let there be no mistake: There is nothing against differences in cases where information about differences is really interesting, e.g. in the case of selecting the best students).

4 Let us look at objectives of instruction as a solution to the problem. Why not describe the objectives in measurable behavior, construct a scale in which a point not far from the maximum means mastery of the aimed at objectives. Then, when we find the normal distribution of scores, why not change our objectives and criterion in such a manner that next time more pupils will reach a certain minimum. Or: give the students early enough access to the description of criterion-behavior and next time they probably can do better. Now both solutions are not realistic because, when we should do so - lower the criterion level or stimulate the students to work toward the critical level of attainment - we will find a negative skewed frequency distribution of scores and our measurement-specialist will tell us that the scale is worthless. So again it does not help, clarifying or reconsidering the wished for level of achievement, as long as we register the student scores in such a manner that the eye is caught by the differences between student scores.

5 But we can improve instruction and thus solve the problem. We can take the position of the experimental psychologist and adapt our instruction to the characteristics of the learner including his entrance knowledge. As John B. Carroll (1963) suggests we provide time enough for mastery. As Bloom (1971) proposes, we use a teaching-learning strategy after carefully providing for dividing the course in units and giving self growing devices as formative tests for the pupils. Now far more students will reach a cognitive level formerly attained by a small percentage. But we found in the first grade that, after improvement of primary reading, teachers adapt their frame of reference when at the end of the course they look at the higher score distributions (Van Calcar 1972). They give the same amount or even more insufficient marks. They simply do not believe their eyes.

6 Now why should they believe their eyes? Is it not natural for human beings to structure what is observed? Does not always the description of a group, of a set, of a collection, make use of comparison of the elements of that collection? We give a description of a collectivity and we indicate the status of an individual element of that set both by comparing.
Always when teachers observe score distributions they are more or less psychologically obliged to structure that group of scores in higher, middle and lower and to judge the individual attainment according to the place in the score distributions. And do not blame them.

7 The frequency table of raw scores and the frequency polygoon or frequency curve are no longer usable, when our interest is - as so many of us claim - in the registration of the fact that a student has mastered the content of a course and in recording the mastery level of the whole group. Now why not simply translate these two interests in simple questions and a simple registration? The two questions are: (a) for student x: did he reach the criterion level? and (b) when not 100% of my students reached the criterion how many of them reached each of a number of lower levels?
These are the two simple questions. And let us <u>not</u> talk here about the necessity for the teacher, who is interested in his pupils, to know more about their competence level for all kinds of reasons. In the first place he does know the investment of extra time and attention for student x and in the second place he can make use of the available standardized tests when selection or prediction is concerned. Now in the next two paragraphs I will make two notes, one about reasonable goalsetting and one about the background of this plea.

8 Note: Setting a criterion in class, does not mean in my conception that this criterion will never change. When all efforts did not result in attainment of the desired point on the scale by a considerable proportion of the students, we cannot but conclude that the criterion really was too high for our class.

In that case, we have to be realistic and lower the criterion level, either at the very moment or, later, with our next group, after considerable deliberation or discussion. After all, what is the use of an objective that cannot be reached. (Of course: motivating the students to reach for an ideal, yes. One could say that.)

9 Note: In formulating the two simple questions and the plea for another registration and representation of the performances, I had the following background in mind: We are talking about:
(a) a sequentially build curriculum in which the curriculum maker and the teacher did their utmost best to optimalize teaching-learning circumstances.
(b) the same content, as common part of the curriculum, has to be mastered by all students; though there may be some individualizing, the end examination under consideration is taken by all the students and they really have to do so, before they can go on with the next course.
(c) the school does not expect the student-achievement to take the form of the normal curve in all circumstances.
(d) no external examination covering the same content is to be kept in mind when reformulating the criterion level or the percentage of students ro reach a certain level.

10 There were two tasks for us. One was the recording for student x of his reaching of the critical point on the attainment-scale. The other was the recording of the master level of the whole group.

Now the first, once we do have a point which is defined as criterion score, is easy enough. When the criterion is set as 80% of the maximum raw score (which score defines the content of the course), there is nothing else to register, as an answer to the question if the student did already reach the criterions level, than "yes" or "not yet".

11. Concerning the recording of the performance of the class, one has to give more information. After the answer is given how many students reached the criterion score, the next question is about the number or percentage of students reaching a given lower level and so on. Often also the question is posed after the mastery level of half of the students: which score was reached by 50% of the students? Or any other percentage. Information to such questions can be given in a tabel in which for all registered raw scores is recorded how many students got that score or less. When we convert both registrations, of raw scores and of students, in percentages, we get an easily comparable registration of class mastery on performance tests. What I did for some empirical score distributions, was making a table with 5 columns, the first two for possible raw scores and percentage of maximum (content) score, the third, fourth and fifth for the number reaching a certain score, the diminishing (decumulative) number of students and the percentage of students reaching a certain score or lower. In the graphical presentation (of 100 x 100 millimeter) matching the table (columns II and V) both questions

"How many reached such and such score?" and "How much was mastered by such and such proportion of the students" can be answered.

12 In tables 1, 2 and 3 are given examples of empirical cases; the connected standard mastery curves are given in figures 1, 2 and 3. In explanation of the three figures can be said:
FIGURE 1: As part of a larger school testing program at the end of primary school, all pupils in grade six in Amsterdam took an arithmetic-test of 25 4-choice items and a grammar-test of 35 items. The first two schools of a sample of schools were the numbers 19 and 33. Obviously the mastery level for arithmetic in both schools was lower than for grammar.
FIGURE 2: In the same program, in 1969, a total of 257 4-choice items on language and arithmetic were taken by 10.000 pupils in Amsterdam. Total scores per student were recorded and represented graphically as standard mastery curve. The normal curve (including guessing) also is given in the same figure.
FIGURE 3: In December 1972 in Twente 92 graduate students in engineering took a 56 (4-choice) item-test on Capita Selecta in Education. The left curve gives their mastery level. After one month 13 of the 92 students took part in a parallel test of 48 items. Most of them in December had been below the cutting score of 60%. After replacing their old scores by the new scores, the right curve describes the mastery of the group of 92 students again.

13 The advantage of the registration and graphical presentation, in standard form of 4 inch or 100 mm square, seems me to be:

a) when simple tables for the conversion of raw scores in percentages are used: quick scoring.

b) clarity for both students and teachers about the measure of mastery.

c) attention is drawn to group-mastery and not to score-distribution.

d) clear illustration of surface to be filled right of the curve.

e) standardized visualisation of ideal curves and levels of mastery.

14 After observing the poor results of many achievement tests, "poor" when considered as a description of the content of a course, I have been thinking about explanations for these unsatisfactory state of affairs. In the next paragraphs I will try to find back the non-mastery strategy which must have been applied to produce such curves. Let us take a course of some length, perhaps 100 hours of learning, consisting of ten units of about the same difficulty. The students have to work the sequence in a fixed order. At the end of each unit a test is taken by the students, but the results are kept in the files. At the end of the course a summative test is taken, which is perfectly representative for the content of the whole course. This test really is a parallel test of all ten foregoing tests matching the units. The results of a hypothetical group of students on this summative test are presented hereafter as standard mastery curves.

15 We * designed four different cases, in which the passing percentages per unit for the first trial and in later trials differed. We had in mind both the shape of the empirical curves and the straightforward instructional models with manageable variables, described and used by Carroll (1963) and Bloom (1971). Does course-aptitude of the learner change? We thought not. Does quality of instruction, in the interplay with learnercharacteristics, change during the course? Perhaps it does. Are the affective learner-characteristics and "perseverance" changing during the course? Maybe. In any case, we tried to estimate the value of some variables responsible for lowering succespercentages for dropouts who try again.

16 The four cases, A, B, C and D, described below led to computations of which the outcomes are given in Table 4. The matching standard mastery curves are given in figure 4.
Case A: On the first unit 90% of the students will succeed. No student will succeed on one of the following units, when he did not master the foregoing unit. As the difficulty of all tasks is the same, again and again 90% of the students beginning a task will finish it with succes, provided they succeeded on all the foregoing units. When a indicates the succespercentage and b the number of units, the percentage of succes of the whole group on the b-th unit will be: $R_b = (\frac{a}{100})^b \times 100$. After ten units 35% of the original group will be left.

Case B: When a is 90 and b is 10, the group of students succeeding on all ten units will be 35%. We suppose now, that after failing one or two of the tests, the slow learners (or, in our case: drop outs) get a new chance after repeated or changed instruction. We hypothesize that of the group who failed the foregoing (b-1) test, (C_1)% will succeed and of the group who failed the two foregoing tests (b-1; b-2), (C_2)% will succeed. Let C_1 be 70 and C_2 : 40. Now the mastery curve gives quite another picture and

*Henk Plessius, Bach. Math. of Twente University of Technology assisted me in the mathematical part of this.

after ten tests even 80% of the orginal group has passed all ten test.

$$R_b = \frac{a}{100} \cdot R_{b-1} + \frac{c_1}{100} \cdot \frac{100-a}{100} \cdot R_{b-1} + \frac{c_2}{100} \cdot \frac{100-c_1}{100} \cdot \frac{100-a}{100}$$

Case C: The difficulty level is 80 for all the units. In this case, c_1 and c_2 both change as a function of b. We take $c_1 = 100 - 10b$ and $c_2 = 100 - 20b$. As said, a = 80. The resulting mastery curve looks more like the empirical curves than the curves A and B: the surface of the square covered is more real than in the case of B. The scape of the curve C is more like reality than curve A. We tried to continue both tendencies in the next case.

Case D: As the influence of c_2 was rather small in these and other cases that we calculated, we take only one c. We take a = 50 and $c = f(b) = 100 - 2b^2$. Of course this choice is not by accident. The resulting curve D now is much like the normal curve. So, when the difficulty or learnability of the tasks is about 50% and the relearning-rate after initial failing on a task is diminishing rapidly as the course goes on, the result can be a mastery curve like we mostly do find in reality.

17 The above four cases of non-mastery strategies may serve as illustrations of theorizing about the causes of the shape of empirical mastery curves. I am convinced of the weaknesses of these first trials but also it is my conviction that this type of theorizing could help us to show what actually happens when a group of students is learning. My hope is that, after observing the differences between what should be the result of learning and the real results (including a simple theory about the causes) we shall be apt to do something about it, more than before. There are two final remarks, one about theorizing in the case of mastery-strategies and the other about biology and skew curves.

18 The four described cases all presuppose that the achievement testscores are kept hidden for students and teachers. Mastery on any task also seems to be perfect mastery of the content which later is represented in the summative test. Now, in practical circumstances the score on the formative test is known and mastery on this test can be defined as a percentage of the maximum raw score. In such cases, we have to reckon with the tendency of students to develop a sensible strategy for applying their energy. Van Naerssen (1970) showed, that students are strategics. We studied two introductory courses in our University, where the prescribed mastery level on intermediate tests was 70% and 80% before the student could go on with the next task. In figure 5 we give (without a table) the empirical mastery curves for the course on "Materiaalkunde" (M) and "wiskunde" (W). For M the prescribed mastery level was 80%, for W it was 70%.

As explanations for the M-curve we give two different possibilities, both of which hold that the succesratio of the slow or failing students is going down so far, that c takes a negative value.
For the time being we refrain from an explanation.
The two possibilities are:

$a_b = 100 - 5b$	$a_b = 70$
$c_b = 100 - 2b$ for $b \leqslant 7$	$c_b = 100 - 1,5b$ for $b \leqslant 7$
$c_8 = 0$	$c_8 = -50$
$c_9 = c_{10} = -100$	$c_9 = c_{10} = -200$

19 Kapteyn (1903) from the astronomic laboratory in Groningen, Netherlands wrote a paper in which he stated that causes independent of the size of the individuals, in biology, seem to produce normal curves, and causes dependent on that size produce skew curves. And he supposes the latter case to be the more general one in nature. Now when we substitute "quality of instruction", and "time allotted to the student" for causes dependent of the individual and we substitute "aptitude" or "entry behavior" for "size" of the individual, there is an analogy. The difference with biology is, that in the teaching and learning situation we can do something about the "size" of the individual and so we can consciously produce skew curves.

20 With our modern technical equipment we can generate more conveniently than in the days of Galton all kinds of curves. Kapteyn (1903) devised a machine (like Galton's apparatus for the normal curve) which illustrated the genesis of the frequency curve for the particular case that the effect of the causes of grow is proportional to the characteristics of the individual. We could do the same.
And in the case that other sciences already know enough about skew curves, why not define ideal skew curves, mastery curves on the bases of different percentages, and staninescores that do not force the teacher to strive after normal score distributions May be we need other appendices in the statistical handbooks for education.

++++++++++++++++

FIGURE 1

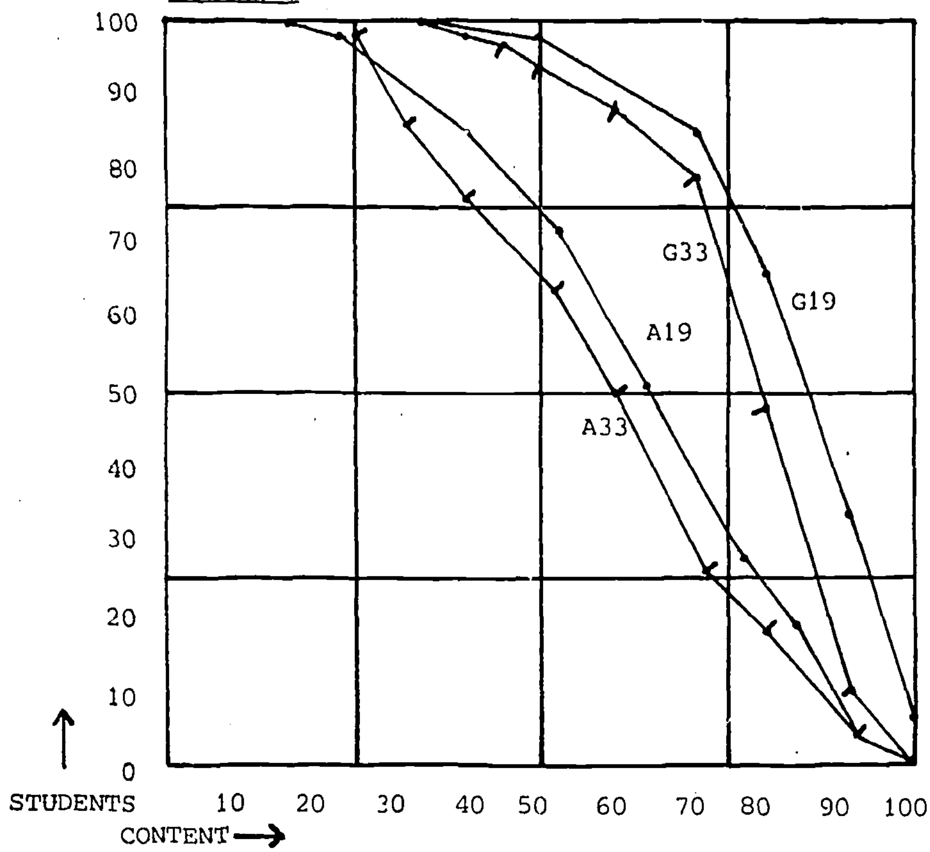


FIGURE 2

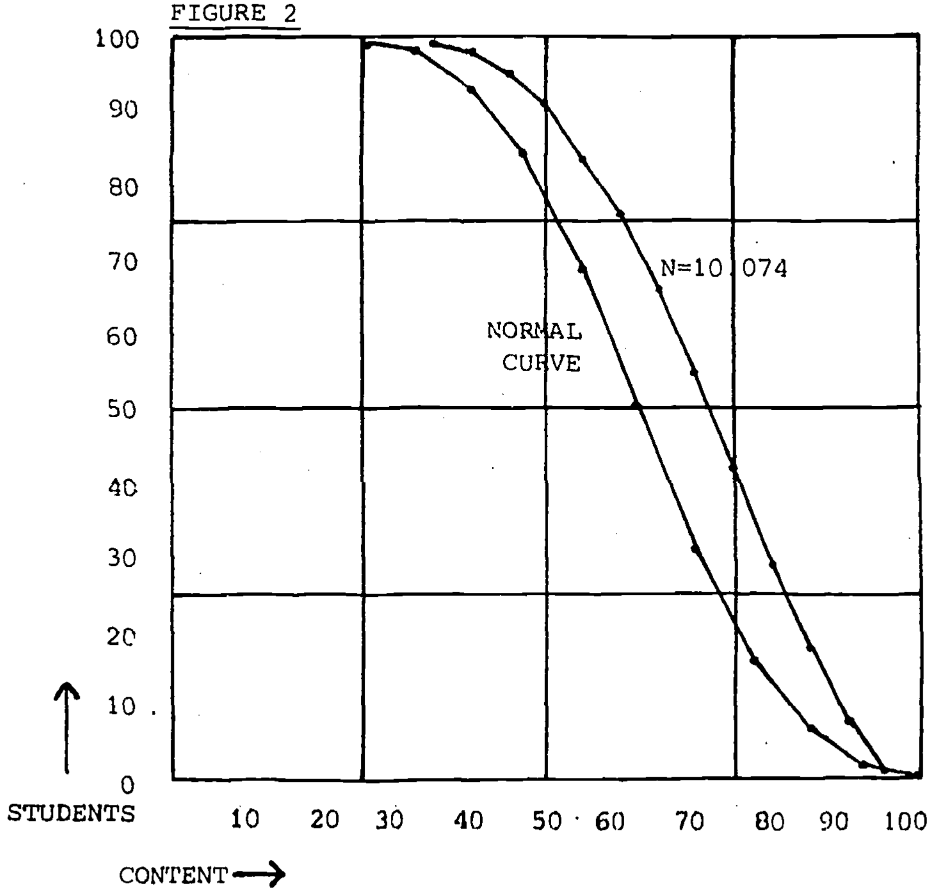


FIGURE 3

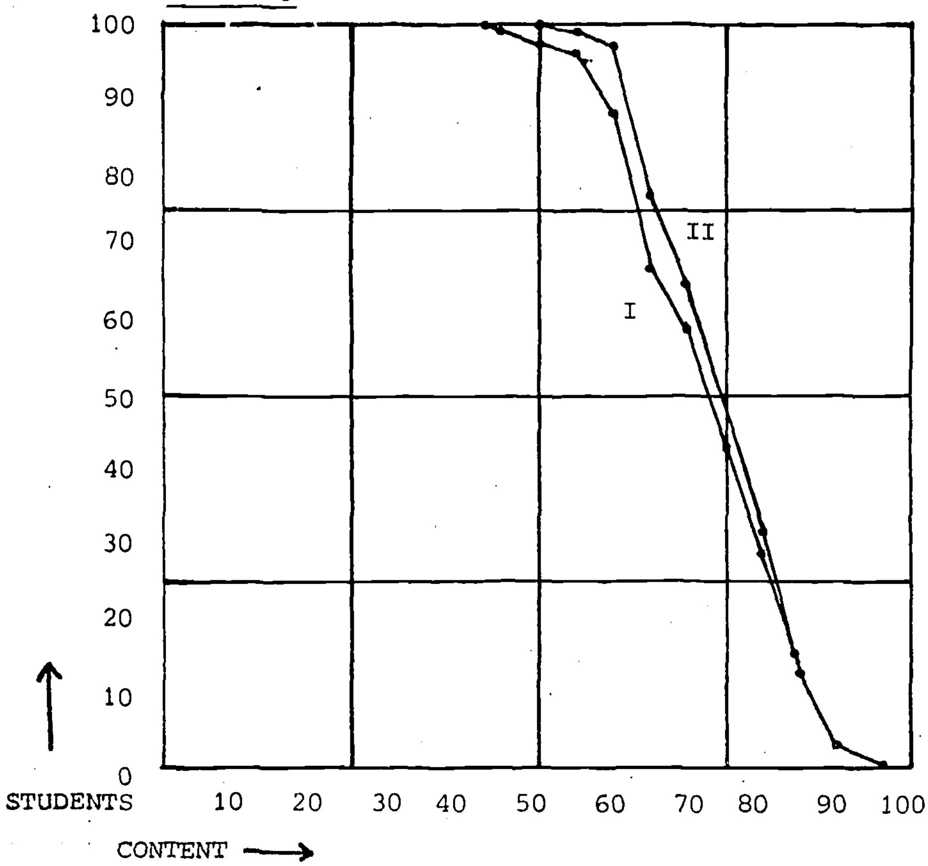


FIGURE 4

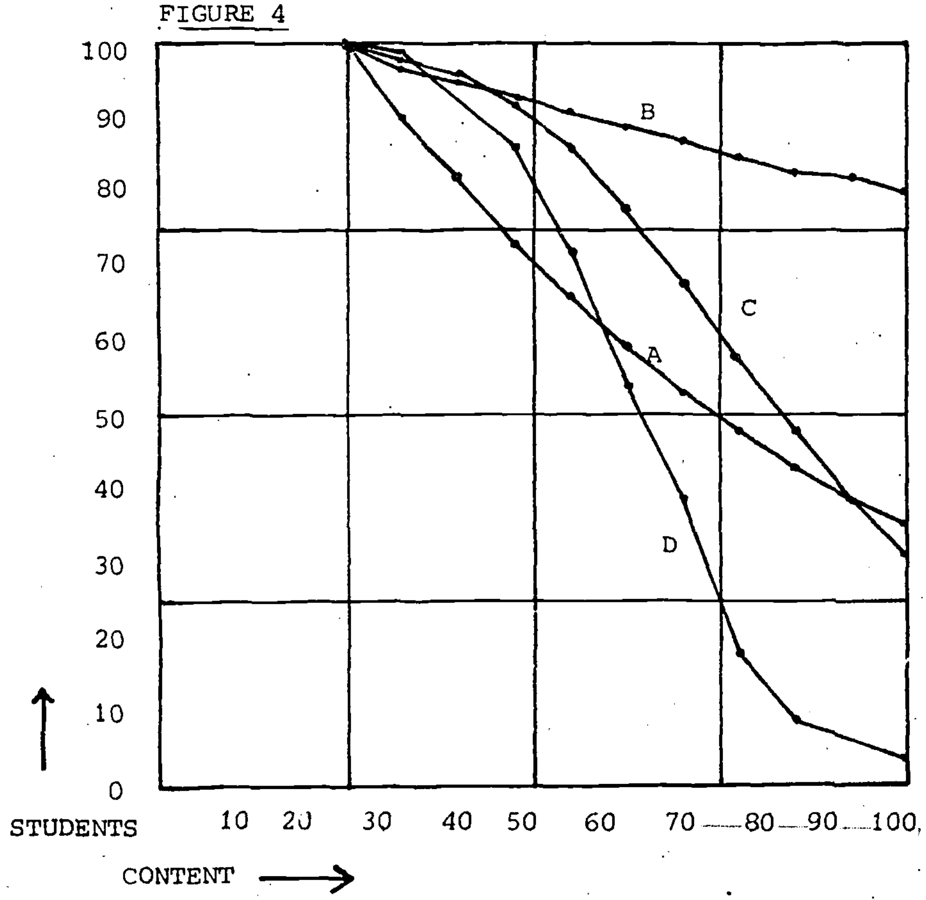


FIGURE 5

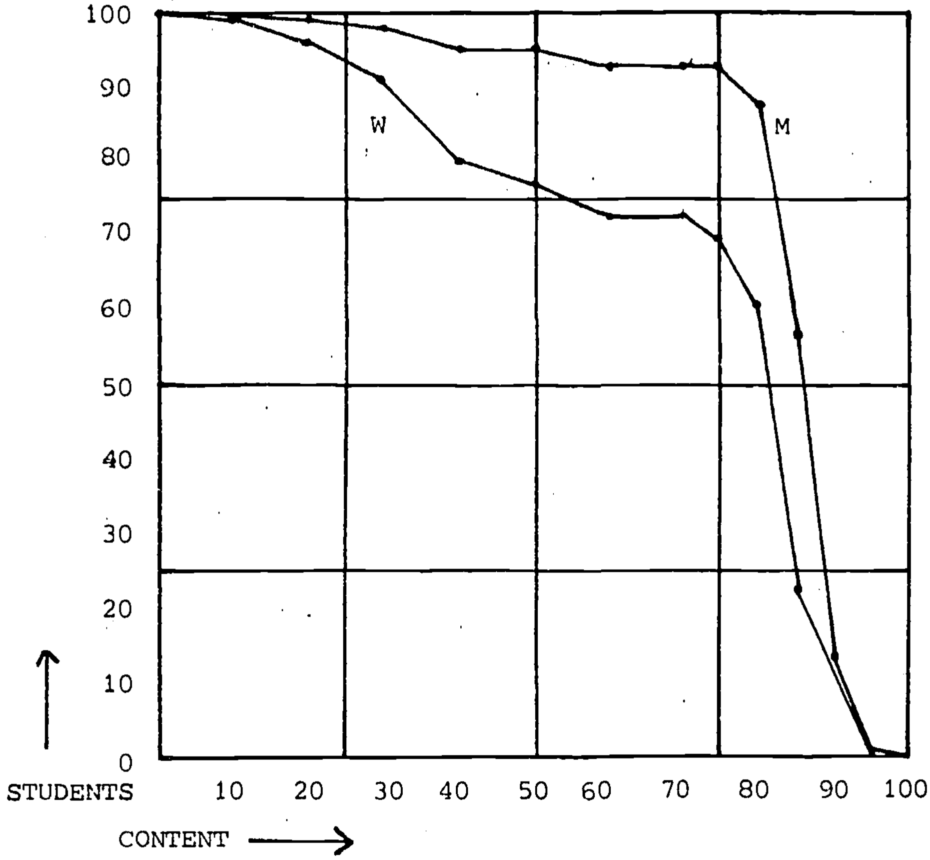

100
90
80
70
60
50
40
30
20
10
0
STUDENTS
W
M
10
20
30
40
50
60
70
80
90
100
CONTENT

TABLE 1

ARITHMETIC SCHOOL 19 CONT. STUDENTS					ARITHMETIC SCHOOL 33 STUDENTS			GRAMMAR SCHOOL 19 CONT. STUDENTS					GRAMMAR SCHOOL 33 STUDENTS		
I	II	III	IV	V	III	IV	V	I	II	III	IV	V	III	IV	V
	%			%			%		%			%			%
4	16	1	53	100	3	50	100	12	34	1	53	100	1	50	100
5	20	0	52	98	1	47	98	13	37	0	52	98	2	49	98
6	24	1	52	98	0	46	92	14	40	0	52	98	0	47	94
7	28	3	51	96	3	46	92	15	43	0	52	98	0	47	94
8	32	3	48	91	4	43	86	16	46	0	52	98	0	47	94
9	36	0	45	85	1	39	78	17	49	0	52	98	0	47	94
10	40	5	45	85	2	38	76	18	52	2	52	98	3	47	94
11	44	1	40	75	2	36	72	19	54	0	50	94	0	44	88
12	48	1	39	74	2	34	68	20	57	3	50	94	0	44	88
13	52	4	38	72	3	32	64	21	60	0	47	89	0	44	88
14	56	5	34	64	4	29	58	22	63	0	47	89	2	44	88
15	60	2	29	54	3	25	50	23	66	1	47	89	0	42	84
16	64	4	27	51	5	21	42	24	69	1	46	87	5	42	84
17	68	4	23	43	3	16	32	25	71	1	45	85	2	37	74
18	72	4	19	36	1	13	26	26	74	6	44	83	4	35	70
19	76	2	15	28	3	12	24	27	77	3	38	72	7	31	62
20	80	3	13	25	3	9	18	28	80	0	35	66	4	24	48
21	84	3	10	19	2	6	12	29	83	5	35	66	4	20	40
22	88	5	7	13	2	4	8	30	86	3	30	57	5	16	32
23	92	0	2	4	1	2	4	31	89	9	27	51	6	11	22
24	96	1	2	4	1	1	2	32	91	3	18	34	5	5	10
25	100	1	1	2	0	0	0	33	94	7	15	28	0	0	0
								34	97	5	8	15	0	0	0
								35	100	3	3	6	0	0	0

TABLE 2

6TH GRADE TEST				NORMAL CURVE	
CONTENT		STUDENTS		C	S
I	II	IV	V	II	V
	%		%	%	%
13	5	10074	100	25	99,3
26	10	10073	99	32,5	97,7
39	15	10073	99	40	93,3
51	20	10073	99	47,5	84,1
64	25	10073	99	55	69,1
77	30	10065	99	62,5	50,0
90	35	10010	99	70	30,9
103	40	9878	98	77,5	15,9
116	45	9584	95	85	6,7
129	50	9118	91	92,5	2,3
141	55	8441	84	100	0,6
154	60	7630	76		
167	65	6650	66		
180	70	5527	55		
193	75	4238	42		
206	80	2929	29		
218	85	1822	18		
231	90	763	8		
244	95	126	1		
257	100	0	0		

TABELE 3

CONTENT I	II %	STUDENTS I III	IV	V %	STUDENTS II III	IV	V %
24	43	0	92	100	0	92	100
25	45	1	92	100	0	92	-
26	46	1	91	99	0	92	-
27	48	1	90	98	0	92	-
28	50	0	89	97	0	92	-
29	52	0	89	97	0	92	-
30	54	0	89	97	0	92	-
31	55	1	89	97	1	92	100
32	57	4	88	96	2	91	99
33	59	2	84	91	0	89	97
34	61	1	82	89	0	89	-
35	63	3	81	88	0	89	-
36	64	8	78	85	9	89	97
37	66	8	70	76	9	80	87
38	68	4	62	67	6	71	77
39	70	4	58	63	5	65	71
40	72	4	54	59	3	60	65
41	73	7	50	54	8	57	62
42	75	3	43	47	5	49	53
43	77	3	40	43	3	44	48
44	79	5	37	40	6	41	45
45	80	5	32	35	6	35	38
46	82	5	27	29	6	29	32
47	84	4	22	24	5	23	25
48	86	6	18	20	6	18	20
49	87	6	12	13	6	12	13
50	89	2	6	7	2	6	7
51	91	1	4	4	1	4	4
52	93	1	3	3	1	3	3
53	95	1	2	2	1	2	2
54	96	0	1	1	0	1	1
55	98	1	1	1	1	1	1
56	100	0	0	0	0	0	0

TABLE 4

CONTENT % (II)	STUDENTS % (V)			
	A	B	C	D
25	100	100	100	100
32,5	90	97	98	99
40	81	95	96	95
47,5	73	93	92	86
55	66	91	86	72
62,5	59	89	78	54
70	53	87	68	34
77,5	48	85	58	17
85	43	83	48	9
92,5	39	82	39	4
100	35	80	31	2

REFERENCES

Bloom, B.S., Learning for mastery. In James H. Block (ed.). Mastery learning; theory and practice. New York, 1971.

Bloom, B.S., "Individual differences in school achievement: a vanishing point", Education at Chicago, winter 1971, pp. 4-14.

Calcar, Co van, Personal Communication, Amsterdam 1972.

Carroll, John B., "A model of school learning", Teachers College Record, Vol. 64, 8, May 1963.

Ebel, Robert L., Content Standard Test Scores. Educational and Psychological Measurement XXII, 1, 1962, pp. 15-25.

Glaser, Robert, Instructional Technology and the Measurement of Learning Outcomes: Some Questions. American Psychologist, 18, 8, 1963, pp. 519-521.

Groot, A.D. de, De Kernitem-methode voor de bepaling van de caesuur voldoende/onvoldoende. Paedagogische Studiën, 41, 10, 1964, pp. 425-440.

Kapteyn, J.C., Skew Frequency Curves in Biology and Statistics, Noordhoff Publisher, Groningen 1903 (also: E. Steiger & Co., New York).

Naerssen, R.F. van, Over optimaal studeren en tentamens combineren. Public Lecture. University of Amsterdam, 1970.

Nedelsky, L., Absolute grading standards for objective test. Educational and Psychological Measurement, 14, 1954.

Wijnen, W.H.F.W., Onder of boven de maat. Dissertation, Groningen 1971.